> **ⓘ Metadata**
>
> - Id: EU.AI4T.O1.M4.1.4t
> - Title: 4.1.4 Risks and AI-based decision making
> - Type: text
> - Description: Understand the classification of risks linked to the use of AI systems in decision
> - Subject: Artificial Intelligence for and by Teachers
> - Authors:
>   - AI4T
> - Licence: CC BY 4.0
> - Date: 2022-11-15

RISKS ASSOCIATED TO THE USE OF AI SYSTEMS IN DECISION MAKING
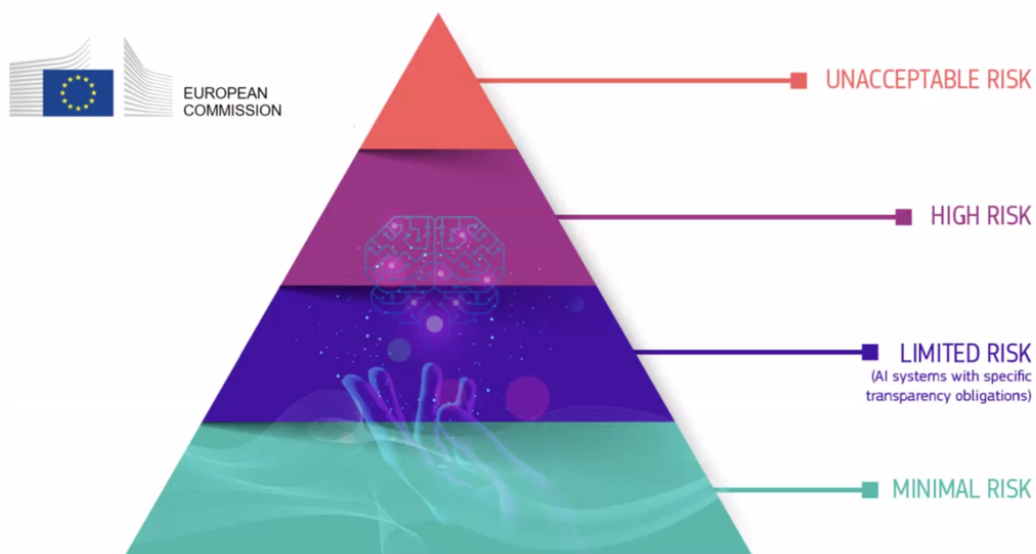
THE 4 LEVELS OF RISK IN AI

The **Regulatory Framework Proposal on Artificial Intelligence**[1] set up by the European Commission (EC) in 2021 identifies four levels of risk linked to the use of AI systems. It is reminded that "*While most AI systems pose limited to no risk and can contribute to solving many societal challenges, certain AI systems create risks that we must address to avoid undesirable outcomes*".
With a focus on the fact that "*it is often not possible to find out why an AI system has made a decision or prediction and taken a particular action. So, it may become difficult to assess whether someone has been unfairly disadvantaged, such as in a hiring decision or in an application for a public benefit scheme*".

Four levels of risk were identified, from minimal to unacceptable:

1. **Unacceptable risk**: All AI systems considered a clear threat to the safety, livelihoods and rights of people will be banned, from social scoring by governments to toys using voice assistance that encourages dangerous behaviour.

2. **High risk**: AI systems identified as high-risk include AI technology used in:
   - Critical infrastructures (e.g. transport), that could put the life and health of citizens at risk;
   - **Educational or vocational training, that may determine the access to education and professional course of someone's life (e.g. scoring of exams)**;
   - Safety components of products (e.g. AI application in robot-assisted surgery);

- Employment, management of workers and access to self-employment (e.g. CV-sorting software for recruitment procedures);
- Essential private and public services (e.g. credit scoring denying citizens opportunity to obtain a loan);
- Law enforcement that may interfere with people's fundamental rights (e.g. evaluation of the reliability of evidence);
- Migration, asylum and border control management (e.g. verification of authenticity of travel documents);
- Administration of justice and democratic processes (e.g. applying the law to a concrete set of facts).

3. **Limited risk**: Limited risk refers to AI systems with specific transparency obligations. When using AI systems such as chatbots, users should be aware that they are interacting with a machine so they can take an informed decision to continue or step back.

4. **Minimal or no risk**: The proposal allows the free use of minimal-risk AI. This includes applications such as AI-enabled video games or spam filters. The vast majority of AI systems currently used in the EU fall into this category.

The proposed 4 levels of Artificial Intelligence risks of the AI Act of the EC.

The classification of educational and vocational training in the high-risk category does not mean that no AI system should be used in those fields but that extra precautions are to be taken. The previously mentioned framework states that "*high-risk AI systems will be subject to strict obligations before they can be placed on the market*".

ETHICS FOR A TRUSTWORTHY AI

AI systems used in education must be trustworthy i.e. complies with the 7 following requirements that AI systems should meet in order to be deemed trustworthy[2]:

- Technical Robustness and safety,

- Privacy and data governance,

- Transparency,

- Diversity, non-discrimination and fairness,

- Societal and environmental well-being,

- Accountability,

- and **Human agency and oversight**: "*AI systems should empower human beings, allowing them to make informed decisions and fostering their fundamental rights. At the same time, proper oversight mechanisms need to be ensured, which can be achieved through human-in-the-loop, human-on-the-loop, and human-in-command approaches*".

---

1. "Laying down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending certain Union Legislatives Acts" - Regulatory Framework Proposal on Artificial Intelligence, European Commission - 2021 ↩

2. "Ethics guidelines for trustworthy AI", European Commission, High-Level Expert Group on AI - 2019. ↩